White Paper

Leveraging EMC Fully Automated Storage Tiering (FAST) and FAST Cache for SQL Server Enterprise Deployments

Applied Technology

Abstract

This white paper introduces EMC's latest groundbreaking technologies, Fully Automated Storage Tiering (FAST) and FAST Cache, and how users can leverage them with SQL Server database applications. This white paper covers the use cases that were lab-tested and the results from those tests. This white paper also discusses best practices for implementing FAST and FAST Cache with SQL Server database applications.

November 2010

Part Number h8091

Table of Contents

Executive summary

Microsoft SQL Server Online Transaction Processing (OLTP) database applications include mission-critical support capabilities and often have stringent I/O latency requirements. Traditionally, these SQL OLTP databases are deployed on a large number of rotating Fibre Channel spindles yet use only a fraction of available disk capacity, a process known as "short-stroking," to meet low I/O latency requirements. Since the effective capacity utilization of these spindles is very low, the cost of the solution is unnecessarily expensive.

This white paper demonstrates that EMC® Fully Automated Storage Tiering (FAST) and FAST Cache technologies provide viable alternative solutions, with regards to I/O performance and total cost of ownership (TCO), to the traditional short-stroking approach for deploying SQL Server. In a SQL TPC-E[1] benchmark test conducted in EMC engineering labs, a baseline configuration with only Fibre Channel spindles is compared to a FAST and FAST Cache configuration that consists of a tiered storage pool using Fibre Channel and SATA II drives for FAST and Flash drives for FAST Cache. The test results show that FAST Cache using Flash drives plus only one-third of Fibre Channel spindles from the baseline configuration can match the baseline performance. Furthermore, FAST allows SATA II drives to complement Fibre Channel drives, providing a low-cost storage tier for inactive data whereas hot data can be automatically promoted to the Fibre Channel tier for faster response. The combination of FAST and FAST Cache offers a flexible and scalable I/O solution for SQL Server enterprise deployments to meet demanding performance and cost-saving objectives.

The need for FAST Cache

SQL Server performance depends heavily on the I/O subsystem. The ultimate limit of throughput and IOPS for SQL Server data read and write is typically limited by disk I/O. Physical I/O performance can make or break a mission-critical SQL Server database application due to its huge impact on the overall user experience.

Over the last few decades, disk drive technology has not evolved as quickly as CPU technology. The fastest rotating spindle (15k rpm) provides only 180 to 200 random 8K IOPS. To improve I/O performance, storage administrators often resort to short-stroking, adding a large number of drives of which only a fraction of the capacity is used. The disadvantages of short-stroking are its poor space utilization and high operational cost, as well as associated high energy consumption.

EMC FAST Cache technology introduces an extra layer of cache using Flash drives between DRAM cache and rotating spindles. This technology provides very low latencies to frequently accessed data, thus improving overall application response times. The hot data in the database application is automatically identified and cached by this newly created cache layer on Flash drives. Therefore, FAST Cache technology provides an optimal I/O solution to meet SQL Server performance needs.

The need for FAST

In the current economic environment, where shrinking IT budgets are the norm, TCO is more important than ever for companies investing in a SQL Server I/O subsystem. The well-known 80/20 rule applies even to enterprise application data. Enterprise data has a temporal nature; some datasets are very hot or frequently accessed compared to other datasets that are rarely accessed. According to industry analysts, 60 percent to 80 percent of operational database data is inactive, increasing as the size of a database grows.[2] Since $/GB cost of SATA II drives is significantly lower than Fibre Channel drives, and SATA II drives consume up to 96 percent less energy per terabyte than a 15k rpm Fibre Channel drive, SATA II drives are an ideal choice for storing inactive data.

However, even though storage administrators understand the benefits of using SATA II drives for inactive data and faster storage media for hot data, they are faced with a constant challenge to classify and store that

[1] *TPC BENCHMARK E. Standard Specification* by the TPC Transactional Processing Performance Council
[2] *Database Archiving: How to Keep Lots of Data for a Very Long Time*. Jack E. Olson, The MK/OMG Press, Massachusetts, 2009

data in the right storage tier. This classification poses a major challenge to storage administrators because of its complexity and manageability. Data classification tools can be helpful, yet downtime is often incurred when data is moved around.

Customers must first create storage pools composed of mixed drive types, such as Flash, Fibre Channel, and SATA II. The storage objects (LUNs) created inside a storage pool can span across all available tiers within that pool. In addition to spanning across available tiers of storage, EMC FAST technology monitors data temperatures and automatically moves the hot and cold data at sub-LUN granularity to the appropriate available tiers within the pool. The data migration is a nondisruptive, application-agnostic operation that storage administrators can control by defining simple data relocation policies. Therefore, FAST is a valid technology for implementing a cost-saving strategy for dealing with active and inactive SQL Server data.

Introduction

This white paper details deployment of EMC FAST and FAST Cache technologies in typical SQL Server enterprise application environments. This white paper discusses the effectiveness of EMC FAST and FAST Cache. These benefits include improved transactions per minute (TPM), lowered latency, and cost savings. The paper also documents the test configurations and performance results of various tests conducted at EMC engineering labs. The testing procedure involved establishing the baseline by running the database benchmark on pure rotating Fibre Channel drives, comparing these baseline numbers with application metrics after adding FAST and FAST Cache to the storage subsystem, and enabling it on all database LUNs where SQL database data files reside.

Audience

This white paper is intended for storage administrators, database administrators, enterprise application administrators, storage architects, customers, and EMC field personnel who want to understand the implementation of FAST and FAST Cache technologies in SQL enterprise environments to provide cost-effective solutions to business applications.

Applying FAST and FAST Cache with a SQL database

Storage administrators need to thoroughly understand the application workload characteristics before implementing any storage technology. While understanding basic information like workload I/O characteristics is important, it is also vital that users understand the active working set of the workload for right-sizing storage tiers in FAST and FAST Cache capacity. The following sections explain in detail how to determine the application I/O characteristics and application working set.

Application I/O characteristics

The relative improvement from any faster storage technology is dependent on whether the application is bottlenecked by the storage subsystem. Numerous sources describe how to use database performance information to determine whether an application is I/O bound. Generally, most SQL DBAs and storage administrators turn to the I/O subsystem because this solution is simpler than restructuring applications to avoid I/O subsystem bottlenecks. However, thoroughly analyzing database-level performance is necessary before investing time and additional resources (such as deploying more Flash drives) to accelerate the storage subsystem.

Understanding the database working set

This step is critical to maximizing the benefits from FAST Cache. Generally, FAST Cache is a very suitable technology for workloads with small random I/O and relatively small working sets. SQL OLTP databases tend to be random in nature with small working sets compared to the total database size. Normally, databases have datasets with varied I/O patterns because of the following reasons:

- OLTP databases tend to be temporal in nature, as the most recent data is more active than older data. This is generally referred to as the working set of a database.

- The relative importance of data changes from object to object. Some tables may be accessed more often than others.

The number of IOPS per gigabyte size of an object, also known as object intensity, changes quite significantly. The best example is an index compared with a table in a database. The relative IOPS received by a database block occupied by an index object are very high compared to the IOPS received by a database block consumed by a table object.

Benchmark setup

To characterize FAST and FAST Cache with SQL OLTP workloads, the following benchmark setup is used: A single-instance SQL 2008 R2 database was deployed on a CLARiiON® CX4-240 storage array and a Dell R900 server. The hardware configuration is shown in Figure 1.

The storage layout is kept simple in the test environment by assigning disk enclosures to one physical server. The real-world storage layouts can be more complex with volumes belonging to various applications spread throughout the array.

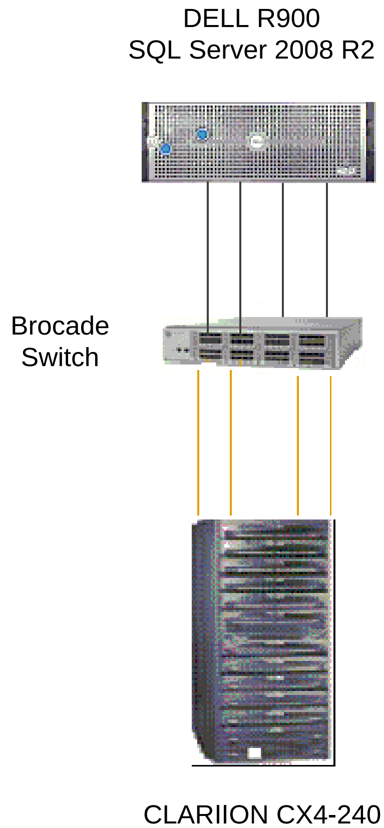

Figure 1. CLARiiON CX4 hardware configuration

Workload description

To demonstrate the effectiveness of the FAST and FAST Cache technologies, EMC engineering conducted SQL TPC-E benchmark testing. The TPC-E database workload simulates an OLTP workload for a brokerage firm with customers who generate transactions related to trades, account inquiries, and market research. The brokerage firm in turn interacts with financial markets to execute orders on behalf of customers and updates relevant account information.

The TPC-E tests are running against SQL Server 2008 R2 with a single instance. The SQL database is configured with a total of 11,000 customers and 50 concurrent users in a TPC-E profile. EMC kept workload characteristics consistent for all use cases.

Use case description

EMC conducted the following tests to understand the benefit of FAST and FAST Cache on SQL OLTP databases and to compare the impact of FAST and FAST Cache on the workload. The goal of these use cases is to show that FAST and FAST Cache can significantly reduce the number of Fibre Channel drives while maintaining the same relative performance once the hot data is cached by FAST Cache.

The test suggests possibilities such as a configuration with fewer rotating Fibre Channel drives or even to replace all Fibre Channel drives with SATA II drives.

The following use cases document this potential. In this experiment, the database data files are the focus of study while the database log file is maintained in a separate LUN without modification for the use cases. *FAST Cache is not enabled on the database log LUN. The logging data I/O pattern is primarily sequential writes, and the storage system memory cache effectively accommodates and coalesces such writes. However, the nature of random read and write patterns exhibited by data LUNs benefits significantly from FAST Cache.*

- **Baseline:** The baseline metric is established on 45 x 600 GB Fibre Channel drives.
- **Use case:** Configure read/write FAST Cache with 4 x 73 GB Flash drives, and two tiers of FAST with 15 x 600 GB Fibre Channel drives and 10 x 2 TB SATA II drives.

Baseline

The baseline was established on an all-Fibre Channel rotating drive configuration as shown in Figure 2. A total of 45 x 600 GB Fibre Channel drives were used for TPC-E data files.

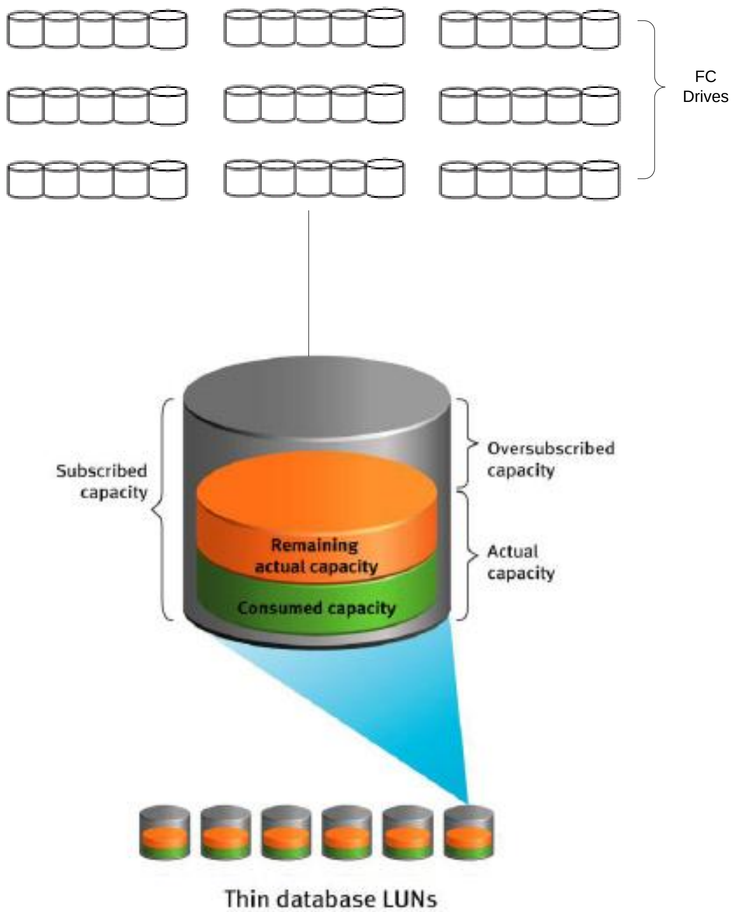

Figure 2. Storage layout for the baseline

CLARiiON LUN layout for the baseline

Thin LUN #1:
- RAID Type: RAID 5
- Drive Type: FC
- Tiering Policy: Manual
- Compression: OFF
- Database files: MSSQL_tpce_root.mdf

Thin LUN #2:
- RAID Type: RAID 5
- Drive Type: FC
- Tiering Policy: Manual
- Compression: OFF
- Database files: Fixed_1.ndf, Growing_1.ndf, Scaling_1.ndf

Thin LUN #3:
- RAID Type: RAID 5
- Drive Type: FC
- Tiering Policy: Manual
- Compression: OFF
- Database files: Fixed_2.ndf, Growing_2.ndf, Scaling_2.ndf

Thin LUN #4:
- RAID Type: RAID 5
- Drive Type: FC

Tiering Policy: Manual
Compression: OFF
Database files: Fixed_3.ndf, Growing_3.ndf, Scaling_3.ndf

Thin LUN #5:
RAID Type: RAID 5
Drive Type: FC
Tiering Policy: Manual
Compression: OFF
Database files: TPCE_Log.ldf

Thin LUN #6:
RAID Type: RAID 5
Drive Type: FC
Tiering Policy: Manual
Compression: OFF
Database files: SQL temp DB

Use case

The purpose of this use case is to illustrate the performance characteristics of FAST and FAST Cache. For that purpose, FAST Cache is configured with four read/write Flash drives and FAST is configured with a Fibre Channel tier and SATA II tier as shown in Figure 3.

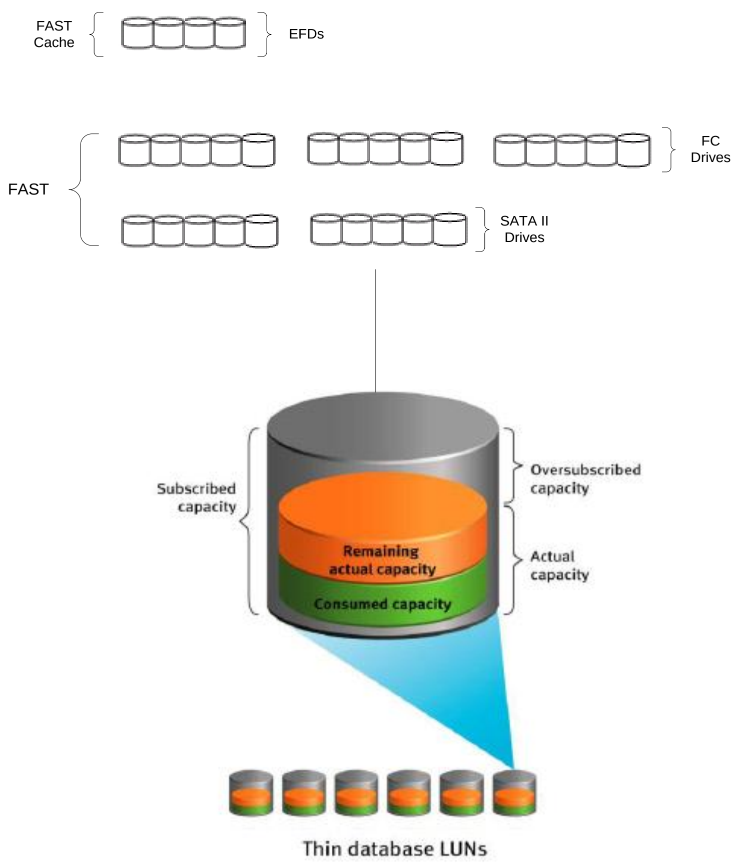

Figure 3. Storage layout for the use case

CLARiiON LUN layout for the use case

Thin LUN #1:
RAID Type: RAID 5
Drive Type: Mixed

Tiering Policy: Auto-tier
Compression: OFF
Database files: MSSQL_tpce_root.mdf

Thin LUN #2:
RAID Type: RAID 5
Drive Type: Mixed
Tiering Policy: Auto-tier
Compression: OFF
Database files: Fixed_1.ndf, Growing_1.ndf, Scaling_1.ndf

Thin LUN #3:
RAID Type: RAID 5
Drive Type: Mixed
Tiering Policy: Auto-tier
Compression: OFF
Database files: Fixed_2.ndf, Growing_2.ndf, Scaling_2.ndf

Thin LUN #4:
RAID Type: RAID 5
Drive Type: Mixed
Tiering Policy: Auto-tier
Compression: OFF
Database files: Fixed_3.ndf, Growing_3.ndf, Scaling_3.ndf

Thin LUN #5:
RAID Type: RAID 5
Drive Type: FC
Tiering Policy: Manual
Compression: OFF
Database files: TPCE_Log.ldf

Thin LUN #6:
RAID Type: RAID 5
Drive Type: FC
Tiering Policy: Manual
Compression: OFF
Database files: SQL temp DB

As shown in Figure 4 on page 11, there is 824.05 GB of data in the SATA II tier that could be moved up to the Fibre Channel tier depending on the data access pattern under the test. Since the Fibre Channel tier is the highest tier in the configuration, the value available to "move up" will always be 0.

Now let's look at Figure 5 on page 11. The data size in the SATA II tier is reduced to 501.03 GB whereas data consumed in the Fibre Channel tier has increased from 363.02 GB to 689.04 GB.

Figure 6 on page 12 shows the FAST auto-tiering schedule policy. Storage administrators can readily adjust the auto-tiering schedule.

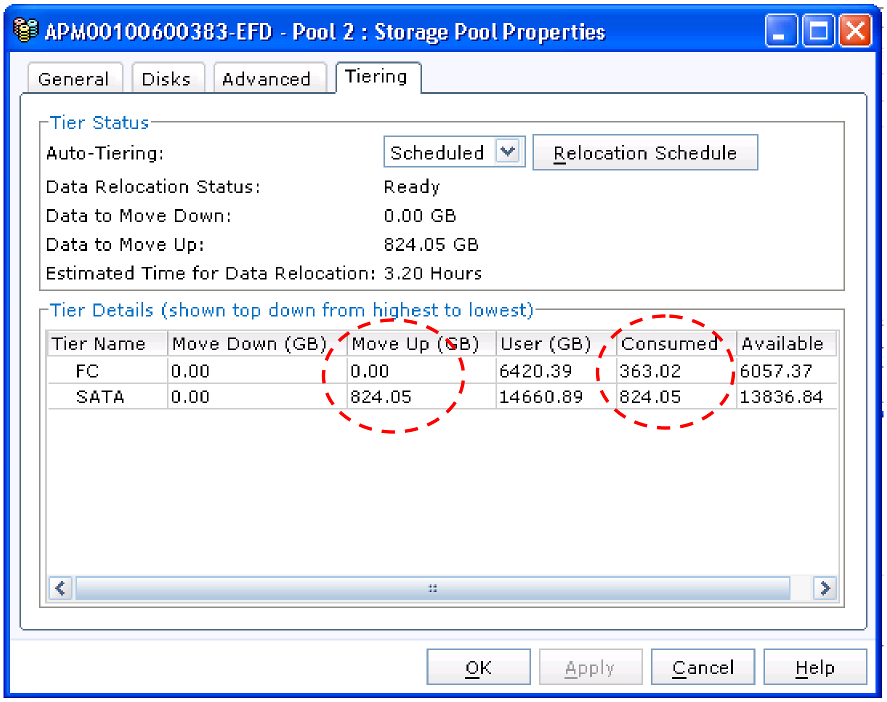

Figure 4. FAST auto-tiering status before the TPC-E run

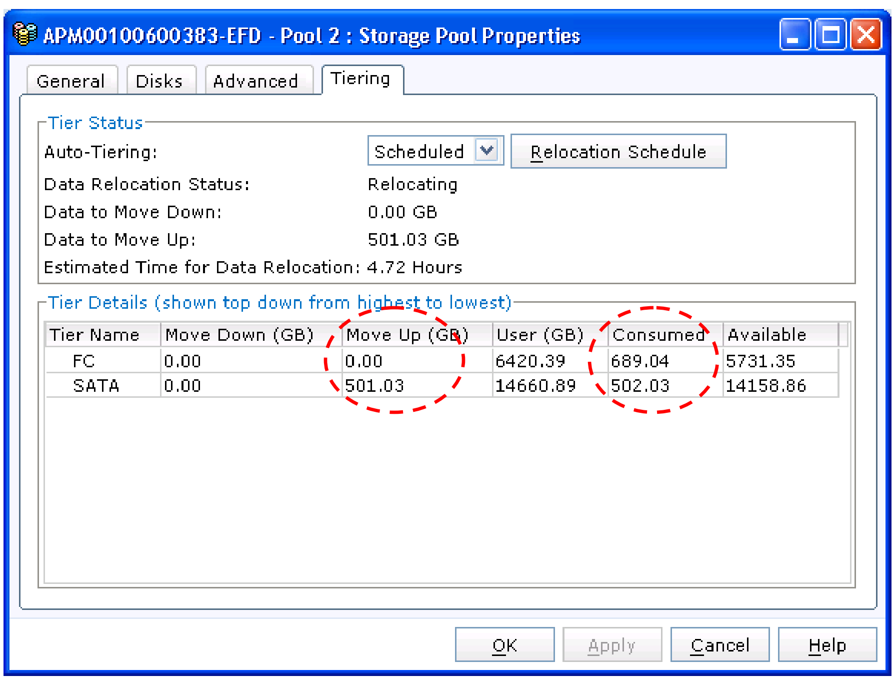

Figure 5. FAST auto-tiering status after the TPC-E run

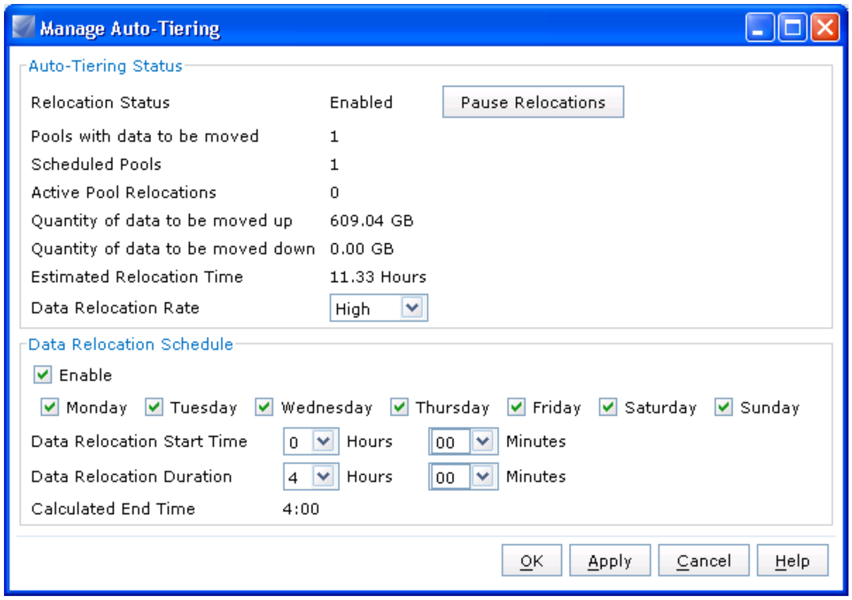

Figure 6. FAST auto-tiering schedule policy settings

FAST Cache can be created through a simple EMC Unisphere™ Analyzer interface as shown in Figure 7. In this example, 4 x 73 GB Flash drives are used to create a read/write FAST Cache in RAID 1/0.

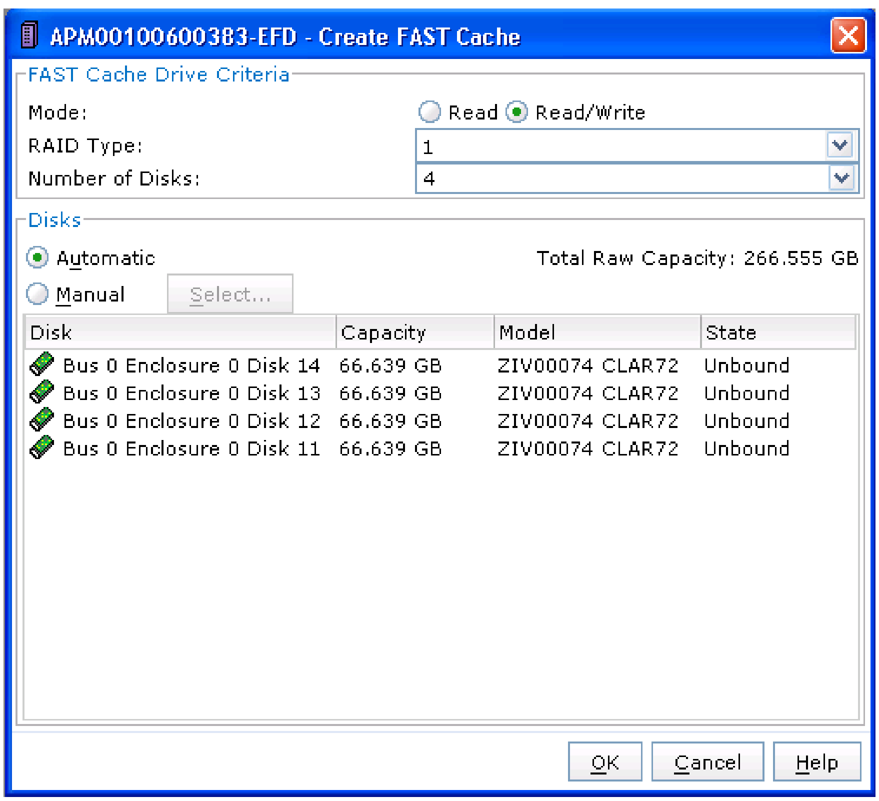

Figure 7. Create FAST Cache

Figure 8 shows the Create FAST Cache properties dialog box after FAST Cache has been created. Notice that the state of FAST Cache is enabled and its size is 133 GB. Initially, the percentages of dirty pages in SP A and SP B are all zero. After the TPC-E run, the percentages of dirty pages for both SP A and SP B become 22 percent. This is the time when FAST Cache will be working for SP A and SP B in the storage system.

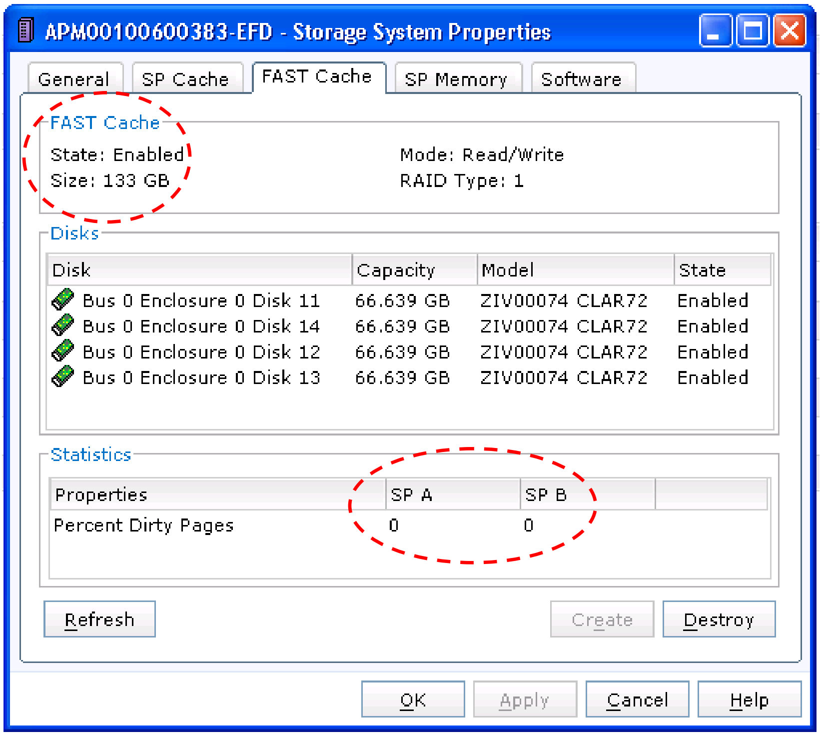

Figure 8. FAST Cache status after creation

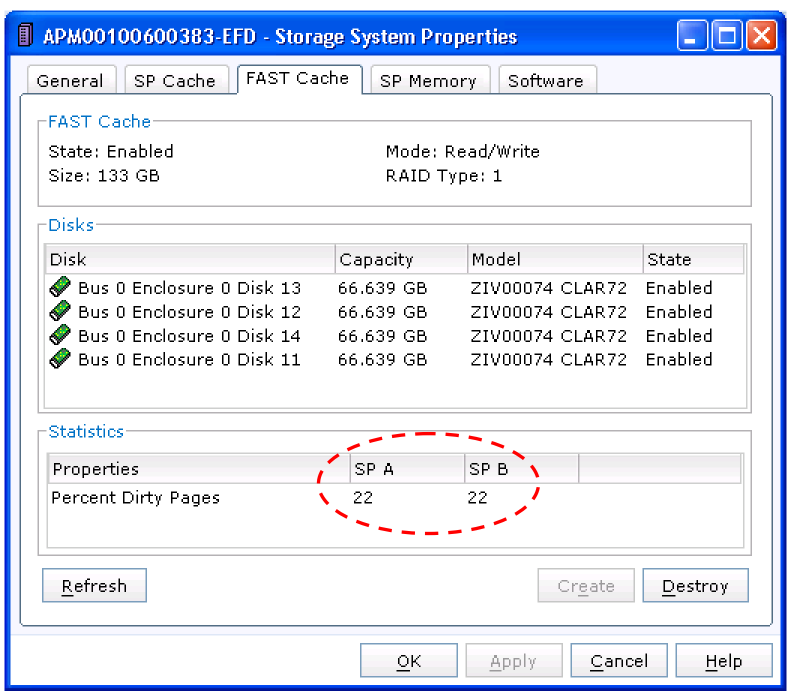

Figure 9. FAST Cache properties page after the TPC-E run

In this configuration, frequently used data is cached by FAST Cache, compensating for the reduced number of Fibre Channel drives and maintaining overall application performance as shown in Figure 10. Analysis reveals that FAST Cache can quickly warm up to match the baseline using 45 Fibre Channel drives. After the warm-up period, FAST Cache not only consistently outperforms the baseline, but also provides the additional benefit of mitigating checkpoint write bursts.

The use case demonstrates that once the hot data gets into FAST Cache, the performance can be achieved with fewer rotating Fibre Channel drives, or SATA II drives.

Reducing the number of rotating Fibre Channel drives does decrease the raw storage capacity. However, because of the implementation of short-stroking to guarantee sufficiently low latencies in real-world deployments, the entire capacity of drives is seldom used. FAST Cache technology enables applications for the first time to leverage the available capacity of these drives. By implementing FAST Cache, rotating media receives only a small fraction of overall I/O and applications will not be impacted significantly by those disk seeks that lower-tier storage must satisfy. This allows consolidation of additional workloads to the same set of available spindles. Costs can be lowered even further through reductions in the power, cooling, and physical space footprint of the data center.

To summarize, the use case configuration has significantly lower TCO in comparison with the baseline while achieving better TPC-E performance. With current EMC pricing for the drives involved in this study, the total costs of disk drives, $/GB, and $/TMP for the use case are reduced by 22 percent, 30 percent, and 28 percent, respectively, in comparison with the baseline.

Note the number of Flash drives required by FAST Cache for equivalent performance relative to the baseline configuration is heavily dependent on the nature of the application and its data access patterns.

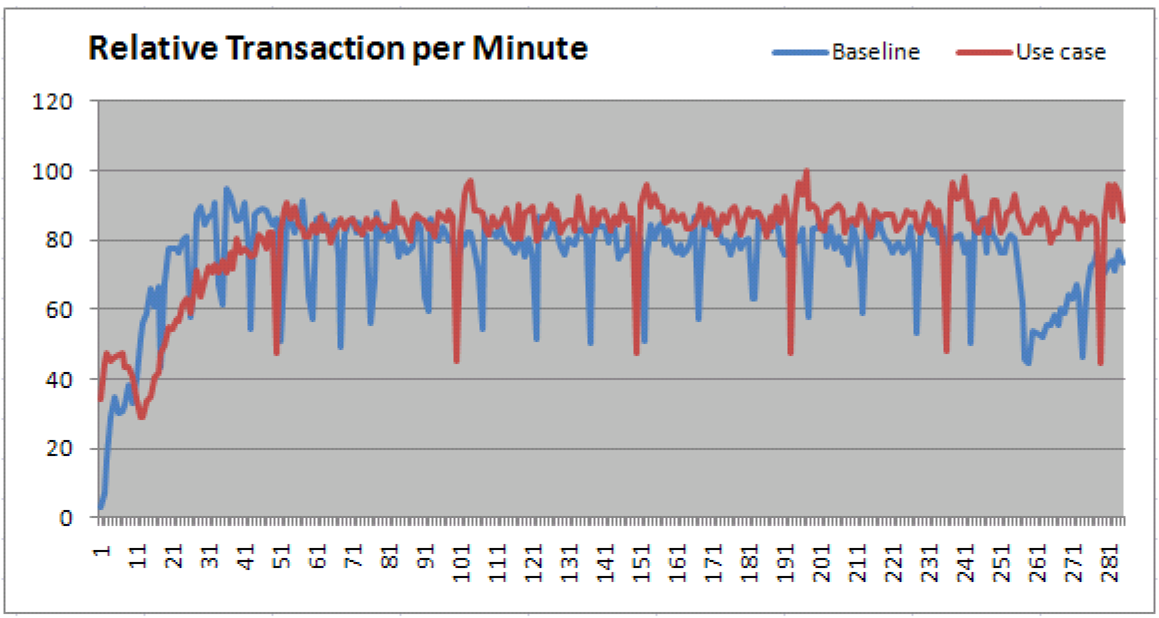

Figure 10. Performance comparison between the baseline and use case

Result summary

The performance results obtained by implementing the FAST and FAST Cache technologies are governed by several factors such as:

- How much the application is currently impacted by the slower I/O subsystem
- Application sensitivity to faster or slower storage
- Application serialization issues

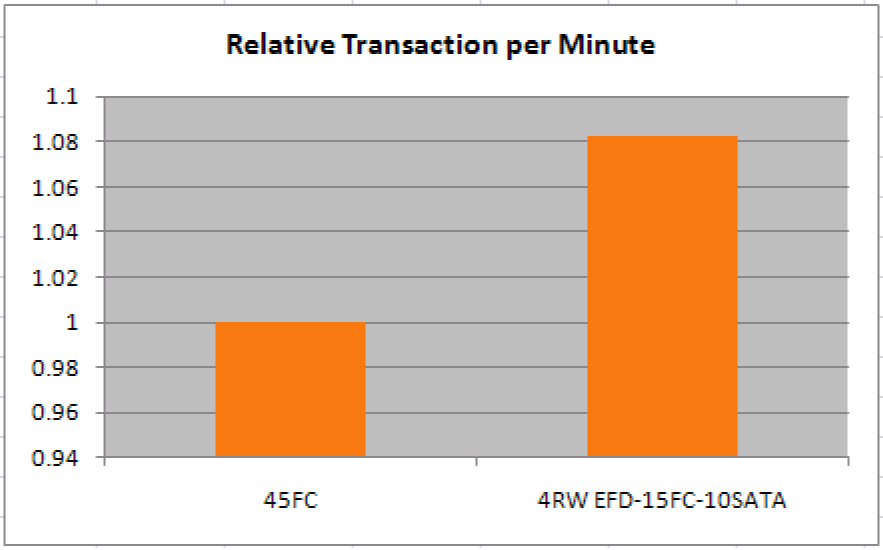

Figure 11. Relative performance comparison against the baseline in TPM with FAST and FAST Cache implementation

Figure 11 shows the relative performance comparison in TPM for TPC-E before and after implementing FAST and FAST Cache. It demonstrates implementation of FAST and FAST Cache technologies permits the I/O subsystem to maintain superior performance by reducing the number of Fibre Channel drives by two-thirds. The performance is made possible by the small investment in Flash drives for FAST Cache, whereas SATA II drives provide for the storage of inactive data drives. Though the use case is a useful frame of reference, remain mindful that TPM ratios are dependent on the nature of application data access patterns.

Conclusion

One of the major advantages of FAST and FAST Cache deployment is to reduce the total spindles required in the I/O subsystem. Most SQL Server enterprise applications have traditionally taken the approach of adding more spindles to meet the demanding requirements for performance. As a result, the I/O subsystem becomes bigger, leading to more energy consumption and more complexity of managing a large number of drives.

The combination of FAST and FAST Cache offers a viable alternative to address the performance challenges of OLTP applications while reducing the total spindles required for the I/O configuration. FAST Cache boosts the performance of the storage system immediately for random access patterns and burst-prone data such as SQL Server checkpoints. On the other hand, FAST allows the storage system to maintain inactive data on low-cost SATA II tier, taking advantage of significant cost and energy savings.

The powerful capabilities of FAST and FAST Cache can improve the TCO of the storage deployment of SQL Server enterprise applications by automatically migrating datasets to appropriate storage tiers. The tests conducted by EMC engineering confirm that, with minimal investment in FAST and FAST Cache, the improved performance, higher capacity utilization, and lowered power and cooling requirements provide customers with a higher return on their storage investment. FAST and FAST Cache are fully automated, completely nondisruptive, and application-agnostic, and therefore can be used with any modern SQL Server enterprise application.

References

The following can be found on EMC.com and Powerlink:

- *EMC CLARiiON CX4 Model 960 Networked Storage System* specification sheet (Powerlink only)
- *Implementing EMC CLARiiON CX4 with Enterprise Flash Drives for Microsoft SQL Server 2008 Databases* white paper
- *Leveraging EMC CLARiiON CX4 with Enterprise Flash Drives for Oracle Database Deployments* white paper
- *Leveraging EMC CLARiiON CX4 with Enterprise Flash Drives for SAP Deployments* white paper
- *EMC CLARiiON Virtual Provisioning – A Detailed Review* white paper
- *EMC* CLARiiON *Best Practices for Performance and Availability: Release 30.0 Firmware Update* white paper